BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 3: e37222
doi: 10.3897/biss.3.37222

OPEN ACCESS

Conference Abstract

Efficient Pipeline for Automating Species ID in new Camera Trap Projects

Sara Beery[‡], Dan Morris[§], Siyu Yang[§], Marcel Simon[§], Arash Norouzzadeh[|], Neel Joshi[§]

‡ Caltech, Pasadena, United States of America
§ Microsoft, Seattle, United States of America
| University of Wyoming, Laramie, United States of America

Corresponding author: Sara Beery (sbeery@caltech.edu)

Received: 12 Jun 2019 | Published: 19 Jun 2019

Citation: Beery S, Morris D, Yang S, Simon M, Norouzzadeh A, Joshi N (2019) Efficient Pipeline for Automating Species ID in new Camera Trap Projects. Biodiversity Information Science and Standards 3: e37222. https://doi.org/10.3897/biss.3.37222

Abstract

Camera traps are heat- or motion-activated cameras placed in the wild to monitor and investigate animal populations and behavior. They are used to locate threatened species, identify important habitats, monitor sites of interest, and analyze wildlife activity patterns. At present, the time required to manually review images severely limits productivity. Additionally, ~70% of camera trap images are empty, due to a high rate of false triggers.

Previous work has shown good results on automated species classification in camera trap data (Norouzzadeh et al. 2018), but further analysis has shown that these results do not generalize to new cameras or new geographic regions (Beery et al. 2018). Additionally, these models will fail to recognize any species they were not trained on. In theory, it is possible to re-train an existing model in order to add missing species, but in practice, this is quite difficult and requires just as much machine learning expertise as training models from scratch. Consequently, very few organizations have successfully deployed machine learning tools for accelerating camera trap image annotation.

We propose a different approach to applying machine learning to camera trap projects, combining a *generalizable detector* with *project-specific classifiers*.

© Beery S et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

We have trained an animal detector that is able to find and localize (but not identify) animals, even species not seen during training, in diverse ecosystems worldwide. See Fig. 1 for examples of the detector run over camera trap data covering a diverse set of regions and species, unseen at training time. By first finding and localizing animals, we are able to:

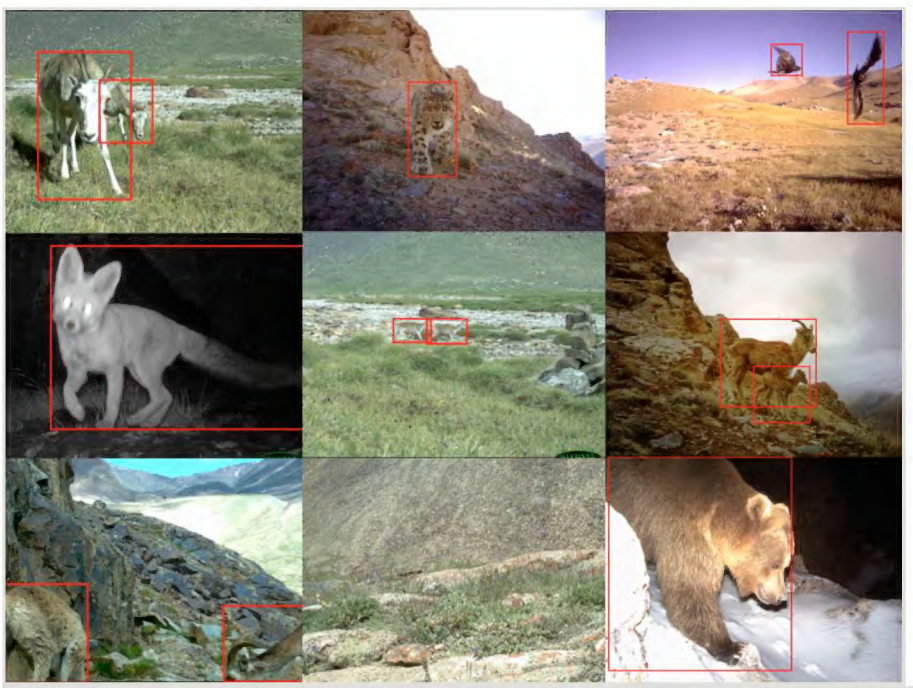

Figure 1.

Animal detection results. Original images provided by the University of Washington.

1. drastically reduce the time spent filtering empty images, and
2. dramatically simplify the process of training species classifiers, because we can crop images to individual animals (and thus classifiers need only worry about animal pixels, not background pixels).

With this detector model as a powerful new tool, we have established a modular pipeline for on-boarding new organizations and building project-specific image processing systems. We break our pipeline into four stages:

1. Data ingestion

First we transfer images to the cloud, either by uploading to a drop point or by mailing an external hard drive. Data comes in a variety of formats; we convert each data set to the CO CO-Camera Traps format, i.e. we create a Javascript Object Notation (JSON) file that encodes the annotations and the image locations within the organization's file structure.

2. Animal detection

We next run our (generic) animal detector on all the images to locate animals. We have developed an infrastructure for efficiently running this detector on millions of images, dividing the load over multiple nodes.

We find that a single detector works for a broad range of regions and species. If the detection results (as validated by the organization) are not sufficiently accurate, it is

possible to collect annotations for a small set of their images and fine-tune the detector. Typically these annotations would be fed back into a new version of the general detector, improving results for subsequent projects.

3. Species classification

Using species labels provided by the organization, we train a (project-specific) classifier on the cropped-out animals.

4. Applying the system to new data

We use the general detector and the project-specific classifier to power tools facilitating accelerated verification and image review, e.g. visualizing the detections, selecting images for review based on model confidence, etc.

The aim of this presentation is to present a new approach to *structuring* camera trap projects, and to formalize discussion around the steps that are required to successfully apply machine learning to camera trap images.

The work we present is available at http://github.com/microsoft/cameratraps, and we welcome new collaborating organizations.

Keywords

camera traps, species classification, computer vision, artificial intelligence, machine learning

Presenting author

Sara Beery

Presented at

Biodiversity_Next 2019

References

- Beery S, Van Horn G, Perona P (2018) Recognition in Terra Incognita. Computer Vision – ECCV 2018 472-489. https://doi.org/10.1007/978-3-030-01270-0_28
- Norouzzadeh MS, Nguyen A, Kosmala M, Swanson A, Palmer M, Packer C, Clune J (2018) Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. Proceedings of the National Academy of Sciences 115 (25). https://doi.org/10.1073/pnas.1719367115